Some New Inequalities for Jeffreys Divergence Measure in Information Theory

SOME NEW INEQUALITIES FOR JEFFREYS DIVERGENCE MEASURE IN INFORMATION THEORY

S.S. DRAGOMIR, J. ŠUNDE, AND C. BUŞE

Abstract. Some new inequalities for the well-known Jeffreys divergence measure in Information Theory are given.

1. Introduction

One of the important issues in many applications of Probability Theory is finding an appropriate measure of *distance* (or *difference* or *discrimination*) between two probability distributions. A number of divergence measures for this purpose have been proposed and extensively studied by Jeffreys [22], Kullback and Leibler [31], Rényi [42], Havrda and Charvat [20], Kapur [25], Sharma and Mittal [44], Burbea and Rao [5], Rao [41], Lin [34], Csiszár [10], Ali and Silvey [1], Vajda [52], Shioya and Da-te [45] and others (see for example [25] and the references therein).

These measures have been applied in a variety of fields such as: anthropology [41], genetics [37], finance, economics, and political science [43], [47], [48], biology [39], the analysis of contingency tables [19], approximation of probability distributions [9], [26], signal processing [23], [24] and pattern recognition [3], [8].

Assume that a set χ and the $\sigma-$finite measure μ are given. Consider the set of all probability densities on μ to be $\Omega := \left\{ p | p : \chi \to \mathbb{R},\, p(x) \geq 0,\ \int_\chi p(x)\, d\mu(x) = 1 \right\}$. The Kullback-Leibler divergence [31] is well known among the information divergences. It is defined as:

$$D_{KL}(p,q) := \int_\chi p(x) \log \left[\frac{p(x)}{q(x)} \right] d\mu(x), \quad p,q \in \Omega, \tag{1.1}$$

where log is to base 2.

In Information Theory and Statistics, various divergences are applied in addition to the Kullback-Leibler divergence. These are the: *variation distance* D_v, *Hellinger distance* D_H [21], χ^2-*divergence* D_{χ^2}, $\alpha-$*divergence* D_α, *Bhattacharyya distance* D_B [4] , *Harmonic distance* D_{Ha}, *Jeffreys distance* D_J [22], *triangular discrimination* D_Δ [49], etc... They are defined as follows:

$$D_v(p,q) := \int_\chi |p(x) - q(x)|\, d\mu(x), \quad p,q \in \Omega; \tag{1.2}$$

$$D_H(p,q) := \int_\chi \left[\sqrt{p(x)} - \sqrt{q(x)} \right]^2 d\mu(x), \quad p,q \in \Omega; \tag{1.3}$$

Date: March 23, 2000.

1991 *Mathematics Subject Classification.* Primary 94Xxx; Secondary 26D15.

Key words and phrases. Jeffreys Divergence, Information Theory, Analytic Inequalities.

The first two authors would like to thank the DSTO for financial support during the work leading to the present paper.

$$D_{\chi^2}(p,q) := \int_{\chi} p(x)\left[\left(\frac{q(x)}{p(x)}\right)^2 - 1\right] d\mu(x), \quad p,q \in \Omega; \tag{1.4}$$

$$D_{\alpha}(p,q) := \frac{4}{1-\alpha^2}\left[1 - \int_{\chi} [p(x)]^{\frac{1-\alpha}{2}} [q(x)]^{\frac{1+\alpha}{2}} d\mu(x)\right], \quad p,q \in \Omega; \tag{1.5}$$

$$D_B(p,q) := \int_{\chi} \sqrt{p(x)q(x)}\, d\mu(x), \quad p,q \in \Omega; \tag{1.6}$$

$$D_{Ha}(p,q) := \int_{\chi} \frac{2p(x)q(x)}{p(x)+q(x)} d\mu(x), \quad p,q \in \Omega; \tag{1.7}$$

$$D_J(p,q) := \int_{\chi} [p(x)-q(x)] \ln\left[\frac{p(x)}{q(x)}\right] d\mu(x), \quad p,q \in \Omega; \tag{1.8}$$

$$D_{\Delta}(p,q) := \int_{\chi} \frac{[p(x)-q(x)]^2}{p(x)+q(x)} d\mu(x), \quad p,q \in \Omega. \tag{1.9}$$

For other divergence measures, see the paper [25] by Kapur or the book on line [46] by Taneja. For a comprehensive collection of preprints available on line, see the RGMIA web site `http://rgmia.vu.edu.au/papersinfth.html`

In [35], Lin and Wong (see also [34]) introduced the following divergence

$$D_{LW}(p,q) := \int_{\chi} p(x) \log\left[\frac{p(x)}{\frac{1}{2}p(x) + \frac{1}{2}q(x)}\right] d\mu(x), \quad p,q \in \Omega. \tag{1.10}$$

In other words, Lin-Wong divergence is represented as follows, using the Kullback-Leibler divergence:

$$D_{LW}(p,q) = D_{KL}\left(p, \frac{1}{2}p + \frac{1}{2}q\right). \tag{1.11}$$

Lin and Wong have shown various inequalities as follows

$$D_{LW}(p,q) \leq \frac{1}{2} D_{KL}(p,q); \tag{1.12}$$

$$D_{LW}(p,q) + D_{LW}(q,p) \leq D_v(p,q) \leq 2; \tag{1.13}$$

$$D_{LW}(p,q) \leq 1. \tag{1.14}$$

In [45], Shioya and Da-te improved $(1.13) - (1.14)$ by showing that

$$D_{LW}(p,q) \leq \frac{1}{2} D_v(p,q) \leq 1. \tag{1.15}$$

For classical and new results in comparing different kinds of divergence measures, see the papers [22]-[45] where further references are given.

In [18], Dragomir and Wang proved, amongst others, the following midpoint inequality

$$\left| f\left(\frac{a+b}{2}\right) - \frac{1}{b-a}\int_a^b f(t)\, dt \right| \leq \frac{1}{4}(b-a)(\Gamma - \gamma), \tag{1.16}$$

provided that f is absolutely continuous and the derivative $f' : [a,b] \to \mathbb{R}$ satisfies the condition

$$\gamma \leq f'(t) \leq \Gamma \text{ for a.e. } t \in [a,b]. \tag{1.17}$$

With the same assumptions for the mapping f, but using a finer argument based in a "pre-Grüss" inequality, the authors of [38] improved (1.1) as follows

$$\left| f\left(\frac{a+b}{2}\right) - \frac{1}{b-a}\int_a^b f(t)\,dt \right| \leq \frac{1}{4\sqrt{3}}(b-a)(\Gamma - \gamma). \tag{1.18}$$

For other results concerning the midpoint and trapezoid inequality, see the recent papers [6]-[7] and the website `http://rgmia.vu.edu.au/`.

The main aim of this paper is to point out some new midpoint and trapezoid type inequalities and apply them for the Jeffreys divergence measure D_J.

2. Some Analytic Inequalities

The following result holds.

Lemma 1. *Let $f : [a,b] \to \mathbb{R}$ be an absolutely continuous mapping on $[a,b]$ with $f' \in L_2[a,b]$. Then we have the inequality:*

$$\left| f\left(\frac{a+b}{2}\right) - \frac{1}{b-a}\int_a^b f(t)\,dt \right| \leq \frac{b-a}{2\sqrt{3}}\left[\frac{1}{b-a}\|f'\|_2^2 - ([f;a,b])^2\right]^{\frac{1}{2}}, \tag{2.1}$$

where

$$[f;a,b] := \frac{f(b)-f(a)}{b-a}.$$

Proof. Start with the following identity which can be easily proved by the integration by parts formula

$$f\left(\frac{a+b}{2}\right) - \frac{1}{b-a}\int_a^b f(t)\,dt = \frac{1}{b-a}\int_a^b m(t) f'(t)\,dt, \tag{2.2}$$

where

$$m(t) := \begin{cases} t-a & \text{if} \quad t \in \left[a, \frac{a+b}{2}\right] \\ t-b & \text{if} \quad t \in \left(\frac{a+b}{2}, b\right] \end{cases}.$$

Using *Korkine's identity*, i.e., we recall it

$$\begin{aligned} &\frac{1}{b-a}\int_a^b u(t)v(t)\,dt - \frac{1}{b-a}\int_a^b u(t)\,dt \cdot \frac{1}{b-a}\int_a^b v(t)\,dt \\ &= \frac{1}{2(b-a)^2}\int_a^b\int_a^b (u(t)-u(s))(v(t)-v(s))\,dtds, \end{aligned} \tag{2.3}$$

and this identity can be proved by direct computation, we may write that

$$\begin{aligned} &\frac{1}{b-a}\int_a^b m(t) f'(t)\,dt - \frac{1}{b-a}\int_a^b m(t)\,dt \cdot \frac{1}{b-a}\int_a^b f'(t)\,dt \\ &= \frac{1}{2(b-a)^2}\int_a^b\int_a^b (m(t)-m(s))(f'(t)-f'(s))\,dtds. \end{aligned} \tag{2.4}$$

However,

$$\int_a^b m(t)\,dt = 0$$

and then, by (2.2) and (2.4), we have the representation:

$$\begin{aligned} &f\left(\frac{a+b}{2}\right) - \frac{1}{b-a}\int_a^b f(t)\,dt \\ = &\frac{1}{2(b-a)^2}\int_a^b\int_a^b (m(t)-m(s))(f'(t)-f'(s))\,dtds. \end{aligned} \tag{2.5}$$

Using the Cauchy-Buniakowski-Schwartz integral inequality for double integrals, we have

$$\begin{aligned} &\frac{1}{2(b-a)^2}\int_a^b\int_a^b |(m(t)-m(s))(f'(t)-f'(s))|\,dtds \\ \leq &\left[\frac{1}{2(b-a)^2}\int_a^b\int_a^b (m(t)-m(s))^2\,dtds\right]^{\frac{1}{2}} \\ &\times\left[\frac{1}{2(b-a)^2}\int_a^b\int_a^b (f'(t)-f'(s))^2\,dtds\right]^{\frac{1}{2}} \end{aligned} \tag{2.6}$$

and as

$$\begin{aligned} &\frac{1}{2(b-a)^2}\int_a^b\int_a^b (m(t)-m(s))^2\,dtds \\ = &\frac{1}{b-a}\int_a^b m^2(t)\,dt - \left(\frac{1}{b-a}\int_a^b m(t)\,dt\right)^2 = \frac{(b-a)^2}{12} \end{aligned}$$

and

$$\begin{aligned} &\frac{1}{2(b-a)^2}\int_a^b\int_a^b (f'(t)-f'(s))^2\,dtds \\ = &\frac{1}{b-a}\int_a^b (f'(t))^2\,dt - \left(\frac{1}{b-a}\int_a^b f'(t)\,dt\right)^2, \end{aligned}$$

then, by (2.5) and (2.6), we deduce (2.1). ∎

Remark 1. *For another proof of this inequality, see* [2].

Remark 2. *Taking into account, by the Grüss inequality, we have that*

$$0 \leq \frac{1}{b-a}\|f'\|_2^2 - ([f;a,b])^2 \leq \frac{1}{4}(\Gamma-\gamma), \tag{2.7}$$

then (2.1) is an improvement of (1.18) in the case when $f' \in L_\infty[a,b]$ *and satisfies (1.17).*

Corollary 1. *For any* $a,b>0$*, we have the inequality*

$$0 \leq (b-a)(\ln b - \ln a) - 2\cdot\frac{(b-a)^2}{a+b} \leq \frac{(b-a)^4}{6\sqrt{a^3b^3}}. \tag{2.8}$$

Proof. Choose $f:(0,\infty)\to\mathbb{R}$, $f(x)=\frac{1}{x}$. Then

$$\begin{aligned} f\left(\frac{a+b}{2}\right) &= \frac{2}{a+b}, \\ \frac{1}{b-a}\int_a^b f(t)\,dt &= \frac{\ln b-\ln a}{b-a}, \\ \frac{1}{b-a}\left\|f'\right\|_2^2-\left([f;a,b]\right)^2 &= \frac{(b-a)^2}{3a^3b^3}, \end{aligned}$$

and then, by (2.1), we get (by the convexity of f) that

$$0\le \frac{\ln b-\ln a}{b-a}-\frac{2}{a+b}\le \frac{(b-a)^2}{6\sqrt{a^3b^3}},$$

which is clearly equivalent to (2.8). ∎

The following lemma also holds.

Lemma 2. *Let $f:[a,b]\to\mathbb{R}$ be an absolutely continuous mapping on $[a,b]$ with $f'\in L_2[a,b]$. Then we have the inequality:*

$$\left|\frac{f(a)+f(b)}{2}-\frac{1}{b-a}\int_a^b f(t)\,dt\right|\le \frac{b-a}{2\sqrt{3}}\left[\frac{1}{b-a}\left\|f'\right\|_2^2-\left([f;a,b]\right)^2\right]^{\frac{1}{2}}. \tag{2.9}$$

Proof. In the recent paper [17], Dragomir and Mabizela proved the following identity which can be easily verified by direct computation:

$$\begin{aligned} &\frac{f(a)+f(b)}{2}-\frac{1}{b-a}\int_a^b f(t)\,dt \\ &= \frac{1}{2(b-a)^2}\int_a^b\int_a^b \left(f'(t)-f'(s)\right)(t-s)\,dtds. \end{aligned} \tag{2.10}$$

Using (2.10) and the Cauchy-Buniakowski-Schwartz integral inequality for double integrals, we have

$$\begin{aligned} &\left|\frac{f(a)+f(b)}{2}-\frac{1}{b-a}\int_a^b f(t)\,dt\right| \\ &\le \frac{1}{2(b-a)^2}\int_a^b\int_a^b \left|\left(f'(t)-f'(s)\right)(t-s)\right|dtds \\ &\le \left[\frac{1}{2(b-a)^2}\int_a^b\int_a^b \left(f'(t)-f'(s)\right)^2 dtds\right]^{\frac{1}{2}} \\ &\quad\times\left[\frac{1}{2(b-a)^2}\int_a^b\int_a^b (t-s)^2\,dtds\right]^{\frac{1}{2}} \end{aligned} \tag{2.11}$$

and as

$$\begin{aligned} &\frac{1}{2(b-a)^2}\int_a^b\int_a^b \left(f'(t)-f'(s)\right)^2 dtds \\ &= \frac{1}{b-a}\int_a^b \left(f'(t)\right)^2 dt-\left(\frac{1}{b-a}\int_a^b f'(t)\,dt\right)^2, \end{aligned}$$

and

$$\frac{1}{2(b-a)^2}\int_a^b\int_a^b (t-s)^2\,dtds = \frac{1}{b-a}\int_a^b t^2dt - \left(\frac{1}{b-a}\int_a^b tdt\right)^2 = \frac{(b-a)^2}{12},$$

then, from (2.11), we deduce the desired inequality (2.9). ∎

Remark 3. *If we assume that f' satisfies (1.17), then by (2.7), we can deduce the inequality*

$$\left|\frac{f(a)+f(b)}{2} - \frac{1}{b-a}\int_a^b f(t)\,dt\right| \leq \frac{1}{4\sqrt{3}}(b-a)(\Gamma-\gamma), \tag{2.12}$$

which improves a similar result from [38] *with the constant $\frac{1}{4}$.*

The following corollary also holds.

Corollary 2. *For any $a, b > 0$, we have the inequality*

$$0 \leq \frac{a+b}{2ab}(b-a)^2 - (\ln b - \ln a)(b-a) \leq \frac{(b-a)^4}{6\sqrt{a^3b^3}}. \tag{2.13}$$

3. Some New Inequalities for Jeffreys Divergence

The following inequalities involving the Jeffreys divergence are known (see for example the book on line by Taneja [46])

$$D_{Ha}(p,q) \geq \exp\left[-\frac{1}{2}D_J(p,q)\right], \quad p,q\in\Omega, \tag{3.1}$$

$$D_{Ha}(p,q) \geq 1 - \frac{1}{4}D_J(p,q), \quad p,q\in\Omega \tag{3.2}$$

and

$$D_J(p,q) \geq 4[1 - D_B(p,q)], \quad p,q\in\Omega, \tag{3.3}$$

where $D_{Ha}(\cdot,\cdot)$ is the Harmonic distance and $D_B(\cdot,\cdot)$ is the Bhattacharyya distance.

The following result holds.

Theorem 1. *We have the inequality*

$$2D_\Delta(p,q) \leq D_J(p,q) \leq \frac{1}{2}\left[D_{\chi^2}(p,q) + D_{\chi^2}(q,p)\right], \quad p,q\in\Omega, \tag{3.4}$$

where D_{χ^2} is the chi-square distance and D_Δ is the triangular discrimination.

Proof. We use the celebrated Hermite-Hadamard inequality for convex functions

$$f\left(\frac{a+b}{2}\right) \leq \frac{1}{b-a}\int_a^b f(t)\,dt \leq \frac{f(a)+f(b)}{2} \tag{3.5}$$

and choose $f(t) = \frac{1}{t}$ to get

$$\frac{2}{a+b} \leq \frac{\ln b - \ln a}{b-a} \leq \frac{a+b}{2ab},$$

which is equivalent to

$$\frac{2(b-a)^2}{a+b} \leq (b-a)(\ln b - \ln a) \leq \frac{a+b}{2ab}(b-a)^2. \tag{3.6}$$

If we choose in (3.6) $b = q(x)$, $a = q(x)$, $x \in \chi$, then we get

$$\begin{aligned} \frac{2(q(x)-p(x))^2}{p(x)+q(x)} &\leq (q(x)-p(x))(\ln q(x)-\ln p(x)) \\ &\leq \frac{p(x)+q(x)}{2p(x)q(x)}(q(x)-p(x))^2 \end{aligned}$$

and integrating over x on χ, we deduce

$$\begin{aligned} 2D_{\Delta}(p,q) &\leq D_J(p,q) \\ &\leq \frac{1}{2}\left[\int_{\chi}\frac{(q(x)-p(x))^2}{q(x)}d\mu(x)+\int_{\chi}\frac{(q(x)-p(x))^2}{p(x)}d\mu(x)\right] \\ &= \frac{1}{2}\left[\int_{\chi}\frac{q^2(x)-2p(x)q(x)+p^2(x)}{q(x)}d\mu(x)\right. \\ &\quad \left.+\int_{\chi}\frac{q^2(x)-2p(x)q(x)+p^2(x)}{p(x)}d\mu(x)\right] \\ &= \frac{1}{2}\left[\int_{\chi}\frac{p^2(x)}{q(x)}d\mu(x)-1+\int_{\chi}\frac{q^2(x)}{p(x)}d\mu(x)-1\right] \\ &= \frac{1}{2}\left[D_{\chi^2}(q,p)+D_{\chi^2}(p,q)\right] \end{aligned}$$

and the inequality (3.4) is deduced. ∎

Using the analytic inequalities established in Section 2, we can prove the following counterpart results as well.

Theorem 2. *For all $p,q \in \Omega$, we have*

$$0 \leq D_J(p,q) - 2D_{\Delta}(p,q) \leq \frac{1}{6}D_*(p,q), \tag{3.7}$$

where

$$D_*(p,q) := \int_{\chi}\frac{(p(x)-q(x))^4}{\sqrt{p^3(x)q^3(x)}}d\mu(x).$$

The proof follows by the inequality (2.8) by a similar procedure as in the proof of Theorem 1 and we omit the details.

By the use of the analytic inequality (2.13), we may state the following theorem.

Theorem 3. *For each $p,q \in \Omega$, we have*

$$0 \leq \frac{1}{2}\left[D_{\chi^2}(p,q)+D_{\chi^2}(q,p)\right] - D_J(p,q) \leq \frac{1}{6}D_*(p,q). \tag{3.8}$$

References

[1] S.M. ALI and S.D. SILVEY, A general class of coefficients of divergence of one distribution from another, *J. Roy. Statist. Soc. Sec B,* **28** (1966), 131-142.

[2] N.S. BARNETT, S.S. DRAGOMIR and A. SOFO, Better bounds for an inequality of the Ostrowski type with applications, *RGMIA Research Report Collection,* **3** (1) (2000), Article 1.

[3] M. BETH BASSAT, f-entropies, probability of error and feature selection, *Inform. Control*, **39** (1978), 227-242.

[4] A. BHATTACHARYYA, On a measure of divergence between two statistical populations defined by their probability distributions, *Bull. Calcutta Math. Soc.,* **35** (1943), 99-109.

[5] I. BURBEA and C.R. RAO, On the convexity of some divergence measures based on entropy function, *IEEE Trans. Inf. Th.,* **28** (3) (1982), 489-495.
[6] P. CERONE and S.S. DRAGOMIR, Midpoint type rules from an inequalities point of view, *RGMIA Research Report Collect.,* **2** (1999), No. 7, Article 1 (`http://rgmia.vu.edu.au/v2n7.html`)
[7] P. CERONE and S.S. DRAGOMIR, Trapezoidal type rules from an inequalities point of view, *RGMIA Research Report Collect.,* **2** (1999), No. 6, Article 8 (`http://rgmia.vu.edu.au/v2n6.html`)
[8] C.H. CHEN, *Statistical Pattern Recognition,* Rocelle Park, New York, Hoyderc Book Co., 1973.
[9] C.K. CHOW and C.N. LIN, Approximating discrete probability distributions with dependence trees, *IEEE Trans. Inf. Th.,* **14** (3) (1968), 462-467.
[10] I. CSISZÁR, Information-type measures of difference of probability distributions and indirect observations, *Studia Math. Hungarica,* **2** (1967), 299-318.
[11] I. CSISZÁR, On topological properties of $f-$divergences, *Studia Math. Hungarica,* **2** (1967), 329-339.
[12] I. CSISZÁR, A note on Jensen's inequality, *Studia Sci. Math. Hung.,* **1** (1966), 185-188.
[13] I. CSISZÁR, On topological properties of $f-$divergences, *Studia Math. Hungarica,* **2** (1967), 329-339.
[14] I. CSISZÁR, A note on Jensen's inequality, *Studia Sci. Math. Hung.,* **1** (1966), 185-188.
[15] I. CSISZÁR and J. KÖRNER, *Information Theory: Coding Theorem for Discrete Memoryless Systems,* Academic Press, New York, 1981.
[16] D. DACUNHA-CASTELLE, *Ecole d′ete de Probability de Saint-Flour*, III-1977, Berlin, Heidelberg: Springer 1978.
[17] S.S. DRAGOMIR and S. MABIZELA, Some error estimates in the trapezoidal quadrature rule, *RGMIA Research Report Collect.,* **2** (1999), No. 5, Article 6 (`http://rgmia.vu.edu.au/v2n5.html`)
[18] S.S. DRAGOMIR and S. WANG, A new inequality of Ostrowski-Grüss type and its applications to the estimation of error bounds for special means and for some numerical quadrature rules, *Comp. Math. Appl.,* **33** (1997), No. 11, 15-20.
[19] D.V. GOKHALE and S. KULLBACK, *Information in Contingency Tables*, New York, Marcel Dekker, 1978.
[20] J.H. HAVRDA and F. CHARVAT, Quantification method classification process: concept of structural α-entropy, *Kybernetika*, **3** (1967), 30-35.
[21] E. HELLINGER, Neue Bergrüirdung du Theorie quadratisher Formerus von uneudlichvieleu Veränderlicher, *J. für reine and Augeur. Math.,* **36** (1909), 210-271.
[22] H. JEFFREYS, An invariant form for the prior probability in estimating problems, *Proc. Roy. Soc. London*, **186** A (1946), 453-461.
[23] T.T. KADOTA and L.A. SHEPP, On the best finite set of linear observables for discriminating two Gaussian signals, *IEEE Trans. Inf. Th.,* **13** (1967), 288-294.
[24] T. KAILATH, The divergence and Bhattacharyya distance measures in signal selection, *IEEE Trans. Comm. Technology.,* Vol COM-15 (1967), 52-60.
[25] J.N. KAPUR, A comparative assessment of various measures of directed divergence, *Advances in Management Studies*, **3** (1984), 1-16.
[26] D. KAZAKOS and T. COTSIDAS, A decision theory approach to the approximation of discrete probability densities, *IEEE Trans. Perform. Anal. Machine Intell.,* **1** (1980), 61-67.
[27] J.H.B. KEMPERMAN, On the optimum note of transmitting information, *Ann. Math. Statist.,* **40** (1969), 2158-2177.
[28] C. KRAFT, Some conditions for consistency and uniform consistency of statistical procedures, *Univ. of California Pub. in Statistics*, **1** (1955), 125-142.
[29] S. KULLBACK, A lower bound for discrimination information in terms of variation, *IEEE Trans. Inf. Th.,* **13** (1967), 126-127.
[30] S. KULLBACK, Correction to a lower bound for discrimination information in terms of variation, *IEEE Trans. Inf. Th.,* **16** (1970), 771-773.
[31] S. KULLBACK and R.A. LEIBLER, On information and sufficiency, *Ann. Math. Stat.*, **22** (1951), 79-86.
[32] S. KULLBACK and R.A. LEIBLER, On information and sufficiency, *Annals Math. Statist.,* **22** (1951), 79-86.

[33] L. LECAM, *Asymptotic Methods in Statistical Decision Theory,* New York: Springer, 1986.
[34] J. LIN, Divergence measures based on the Shannon entropy, *IEEE Trans. Inf. Th.,* **37** (1) (1991), 145-151.
[35] J. LIN and S.K.M. WONG, A new directed divergence measure and its characterization, *Int. J. General Systems,* **17** (1990), 73-81.
[36] H.P. McKEAN, JR., Speed of approach to equilibrium for Koc's caricature of a Maximilian gas, *Arch. Ration. Mech. Anal.,* **21** (1966), 343-367.
[37] M. MEI, The theory of genetic distance and evaluation of human races, *Japan J. Human Genetics*, **23** (1978), 341-369.
[38] C.E.M. PEARCE, J. PEČARIĆ, N. UJEVIĆ and S. VAROŠANEC, Generaliasations of some inequalities of Ostrowski-Grüss type, *Math. Ineq. & Appl.,* **3** (2000), No. 1, 25-34.
[39] E.C. PIELOU, *Ecological Diversity,* Wiley, New York, 1975.
[40] M.S. PINSKER, Information and Information Stability of Random variables and processes, (in Russian), Moscow: Izv. Akad. Nouk, 1960.
[41] C.R. RAO, Diversity and dissimilarity coefficients: a unified approach, *Theoretic Population Biology*, **21** (1982), 24-43.
[42] A. RÉNYI, On measures of entropy and information, *Proc. Fourth Berkeley Symp. Math. Stat. and Prob., University of California Press*, **1** (1961), 547-561.
[43] A. SEN, *On Economic Inequality,* Oxford University Press, London 1973.
[44] B.D. SHARMA and D.P. MITTAL, New non-additive measures of relative information, *Journ. Comb. Inf. Sys. Sci.,* **2** (4)(1977), 122-132.
[45] H. SHIOYA and T. DA-TE, A generalisation of Lin divergence and the derivative of a new information divergence, *Elec. and Comm. in Japan*, **78** (7) (1995), 37-40.
[46] I. J. TANEJA, *Generalised Information Measures and their Applications* (http://www.mtm.ufsc.br/~taneja/bhtml/bhtml.html).
[47] H. THEIL, *Economics and Information Theory*, North-Holland, Amsterdam, 1967.
[48] H. THEIL, *Statistical Decomposition Analysis*, North-Holland, Amsterdam, 1972.
[49] F. TOPSOE, Some inequalities for information divergence and related measures of discrimination, *Res. Rep. Coll., RGMIA,* **2** (1) (1999), 85-98.
[50] G.T. TOUSSAINT, Sharper lower bounds for discrimination in terms of variation, *IEEE Trans. Inf. Th.,* **21** (1975), 99-100.
[51] I. VAJDA, Note on discrimination information and variation, *IEEE Trans. Inf. Th.,* **16** (1970), 771-773.
[52] I. VAJDA, *Theory of Statistical Inference and Information,* Dordrecht-Boston, Kluwer Academic Publishers, 1989.
[53] V.A. VOLKONSKI and J. A. ROZANOV, Some limit theorems for random function -I, (English Trans.), *Theory Prob. Appl., (*USSR), **4** (1959), 178-197.

School of Communications and Informatics, Victoria University of Technology, PO Box 14428, Melbourne City MC, Victoria 8001, Australia.
E-mail address: sever.dragomir@vu.edu.au
URL: http://rgmia.vu.edu.au/SSDragomirWeb.html

E-mail address: Jadranka.Sunde@dsto.defence.gov.au

Communication Division, DSTO, PO Box 1500, Salsbury, SA 5108, Australia.

E-mail address: buse@hilbert.math.uvt.ro

Department of Mathematics, West Timisoara University, B-dul V. Parvan No. 4, RO-1900, Timisoara, Romania